

OSD

OIPE

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/434,382DATE: 11/17/1999
TIME: 13:52:25

Input Set: I434382.RAW

This Raw Listing contains the General Information Section and up to first 5 pages.

1 <110> APPLICANT: Tavtigian, Sean V.
 2 Teng, David H.F.
 3 Simard, Jacques
 4 Rommens, Johanna M.
 5 Myriad Genetics, Inc.

6 <120> TITLE OF INVENTION: Chromosome 17q-Linked Prostate Cancer Susceptibility
 7 Gene

8 <130> FILE REFERENCE: HPC2 Gene

9 <140> CURRENT APPLICATION NUMBER: US/09/434,382

10 <141> CURRENT FILING DATE: 1999-11-05

11 <150> EARLIER APPLICATION NUMBER: US 60/107,468

12 <151> EARLIER FILING DATE: 1999-11-06

13 <160> NUMBER OF SEQ ID NOS: 213

14 <170> SOFTWARE: PatentIn Ver. 2.0

15 <210> SEQ ID NO 1

16 <211> LENGTH: 2481

17 <212> TYPE: DNA

18 <213> ORGANISM: Homo sapiens

19 <220> FEATURE:

20 <221> NAME/KEY: CDS

21 <222> LOCATION: (1)...(2478)

22 <400> SEQUENCE: 1

atg tgg gcg ctt tgc tcg ctg cgg tcc gcg gcc acc atg	48
Met Trp Ala Leu Cys Ser Leu Leu Arg Ser Ala Ala Gly Arg Thr Met	
1 5 10 15	
tcg cag gga cgc acc ata tcg cag gca ccc gcc cgc cgc gag cgg cgg	96
Ser Gln Gly Arg Thr Ile Ser Gln Ala Pro Ala Arg Arg Glu Arg Pro	
20 25 30	
cgc aag gac ccg ctg cgg cac ctg cgc acg cga gag aag cgc gga ccg	144
Arg Lys Asp Pro Leu Arg His Leu Arg Thr Arg Glu Lys Arg Gly Pro	
35 40 45	
tcg ggg tgc tcc ggc cca aac acc gtg tac ctg cag gtg gtg gca	192
Ser Gly Cys Ser Gly Gly Pro Asn Thr Val Tyr Leu Gln Val Val Ala	
50 55 60	
gcg ggt agc cgg gac tcg ggc gcc gcg ctc tac gtc ttc tcc gag ttc	240
Ala Gly Ser Arg Asp Ser Gly Ala Ala Leu Tyr Val Phe Ser Glu Phe	
65 70 75 80	
aac cgg tat ctc ttc aac tgt gga gaa ggc gtt cag aga ctc atg cag	288
Asn Arg Tyr Leu Phe Asn Cys Gly Glu Gly Val Gln Arg Leu Met Gln	
85 90 95	
gag cac aag tta aag gtt gct cgc ctg gac aac ata ttc ctg aca cga	336
Glu His Lys Leu Lys Val Ala Arg Leu Asp Asn Ile Phe Leu Thr Arg	
100 105 110	
atg cac tgg tct aat gtt ggg ggc tta agt gga atg att ctt act tta	384

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/434,382DATE: 11/17/1999
TIME: 13:52:25

Input Set: I434382.RAW

45	Met His Trp Ser Asn Val Gly Gly Leu Ser Gly Met Ile Leu Thr Leu			
46	115	120	125	
47	aag gaa acc ggg ctt cca aag tgt gta ctt tct gga cct cca caa ctg	432		
48	Lys Glu Thr Gly Leu Pro Lys Cys Val Leu Ser Gly Pro Pro Gln Leu			
49	130	135	140	
50	gaa aaa tac ctc gaa gca atc aaa ata ttt tct ggt cca ttg aaa gga	480		
51	Glu Lys Tyr Leu Glu Ala Ile Lys Ile Phe Ser Gly Pro Leu Lys Gly			
52	145	150	155	160
53	ata gaa ctg gct gtg cgg ccc cac tct gcc cca gaa tac gag gat gaa	528		
54	Ile Glu Leu Ala Val Arg Pro His Ser Ala Pro Glu Tyr Glu Asp Glu			
55	165	170	175	
56	acc atg aca gtt tac cag atc ccc ata cac agt gaa cag agg agg gga	576		
57	Thr Met Thr Val Tyr Gln Ile Pro Ile His Ser Glu Gln Arg Arg Gly			
58	180	185	190	
59	aag cac caa cca tgg cag agt cca gaa agg cct ctc agc agg ctc agt	624		
60	Lys His Gln Pro Trp Gln Ser Pro Glu Arg Pro Leu Ser Arg Leu Ser			
61	195	200	205	
62	cca gag cga tct tca gac tcc gag tcg aat gaa aat gag cca cac ctt	672		
63	Pro Glu Arg Ser Ser Asp Ser Glu Ser Asn Glu Asn Glu Pro His Leu			
64	210	215	220	
65	cca cat ggt gtt agc cag aga aga ggg gtc agg gac tct tcc ctg gtc	720		
66	Pro His Gly Val Ser Gln Arg Arg Gly Val Arg Asp Ser Ser Leu Val			
67	225	230	235	240
68	gta gct ttc atc tgt aag ctt cac tta aag aga gga aac ttc ttg gtg	768		
69	Val Ala Phe Ile Cys Lys Leu His Leu Lys Arg Gly Asn Phe Leu Val			
70	245	250	255	
71	ctc aaa gca aag gag atg ggc ctc cca gtt ggg aca gct gcc atc gct	816		
72	Leu Lys Ala Lys Glu Met Gly Leu Pro Val Gly Thr Ala Ala Ile Ala			
73	260	265	270	
74	ccc atc att gct gtc aag gac ggg aaa agc atc act cat gaa gga	864		
75	Pro Ile Ile Ala Ala Val Lys Asp Gly Lys Ser Ile Thr His Glu Gly			
76	275	280	285	
77	aga gag att ttg gct gaa gag ctg tgt act cct cca gat cct ggt gct	912		
78	Arg Glu Ile Leu Ala Glu Glu Leu Cys Thr Pro Pro Asp Pro Gly Ala			
79	290	295	300	
80	gct ttt gtg gtg gta gaa tgt cca gat gaa agc ttc att caa ccc atc	960		
81	Ala Phe Val Val Glu Cys Pro Asp Glu Ser Phe Ile Gln Pro Ile			
82	305	310	315	320
83	tgt gag aat gcc acc ttt cag agg tac caa gga aag gca gat gcc ccc	1008		
84	Cys Glu Asn Ala Thr Phe Gln Arg Tyr Gln Gly Lys Ala Asp Ala Pro			
85	325	330	335	
86	gtg-gcc-ttg-gtg-gtt-cac-atg-gcc-cca-gca-tct-gtg-ctt-gtg-gac-agc	1056		
87	Val Ala Leu Val Val His Met Ala Pro Ala Ser Val Leu Val Asp Ser			
88	340	345	350	
89	agg tac cag cag tgg atg gag agg ttt ggg cct gac acc cag cac ttg	1104		
90	Arg Tyr Gln Gln Trp Met Glu Arg Phe Gly Pro Asp Thr Gln His Leu			
91	355	360	365	
92	gtc ctg aat gag aac tgt gcc tca gtt cac aac ctt cgc agc cac aag	1152		
93	Val Leu Asn Glu Asn Cys Ala Ser Val His Asn Leu Arg Ser His Lys			
94	370	375	380	

PAGE: 3

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/434,382DATE: 11/17/1999
TIME: 13:52:25

Input Set: I434382.RAW

95	att caa acc cag ctc aac ctc atc cac ccg gac atc ttc ccc ctg ctc	1200
96	Ile Gln Thr Gln Leu Asn Leu Ile His Pro Asp Ile Phe Pro Leu Leu	
97	385 390 395 400	
98	acc agt ttc cgc tgt aag aag gag ggc ccc acc ctc agt gtg ccc atg	1248
99	Thr Ser Phe Arg Cys Lys Lys Glu Gly Pro Thr Leu Ser Val Pro Met	
100	405 410 415	
101	gtt cag ggt gaa tgc ctc ctc aag tac cag ctc cgt ccc agg agg gag	1296
102	Val Gln Gly Glu Cys Leu Leu Lys Tyr Gln Leu Arg Pro Arg Arg Glu	
103	420 425 430	
104	tgg cag agg gat gcc att att act tgc aat cct gag gaa ttc ata gtt	1344
105	Trp Gln Arg Asp Ala Ile Ile Thr Cys Asn Pro Glu Glu Phe Ile Val	
106	435 440 445	
107	gag gcg ctg cag ctt ccc aac ttc cag cag agc gtg cag gag tac agg	1392
108	Glu Ala Leu Gln Leu Pro Asn Phe Gln Gln Ser Val Gln Glu Tyr Arg	
109	450 455 460	
110	agg agt gcg cag gac ggc cca gcc cca gca gag aaa aga agt cag tac	1440
111	Arg Ser Ala Gln Asp Gly Pro Ala Pro Ala Glu Lys Arg Ser Gln Tyr	
112	465 470 475 480	
113	cca gaa atc atc ttc ctt gga aca ggg tct gcc atc ccg atg aag att	1488
114	Pro Glu Ile Ile Phe Leu Gly Thr Gly Ser Ala Ile Pro Met Lys Ile	
115	485 490 495	
116	cga aat gtc agt gcc aca ctt gtc aac ata agc ccc gac acg tct ctg	1536
117	Arg Asn Val Ser Ala Thr Leu Val Asn Ile Ser Pro Asp Thr Ser Leu	
118	500 505 510	
119	cta ctg gac tgt ggt gag ggc aca ttt ggg cag ctg tgc cgt cat tac	1584
120	Leu Leu Asp Cys Gly Glu Gly Thr Phe Gly Gln Leu Cys Arg His Tyr	
121	515 520 525	
122	gga gac cag gtg gac agg gtc ctg ggc acc ctg gct gct gtg ttt gtg	1632
123	Gly Asp Gln Val Asp Arg Val Leu Gly Thr Leu Ala Ala Val Phe Val	
124	530 535 540	
125	tcc cac ctg cac gca gat cac cac acg ggc ttg cca agt atc ttg ctg	1680
126	Ser His Leu His Ala Asp His His Thr Gly Leu Pro Ser Ile Leu Leu	
127	545 550 555 560	
128	cag aga gaa cgc gcc ttg gca tct ttg gga aag ccg ctt cac cct ttg	1728
129	Gln Arg Glu Arg Ala Leu Ala Ser Leu Gly Lys Pro Leu His Pro Leu	
130	565 570 575	
131	ctg gtg gtt gcc ccc aac cag ctc aaa gcc tgg ctc cag cag tac cac	1776
132	Leu Val Val Ala Pro Asn Gln Leu Lys Ala Trp Leu Gln Gln Tyr His	
133	580 585 590	
134	aac cag tgc cag gag gtc ctg cac cac atc agt atg att cct gcc aaa	1824
135	Asn Gln Cys Gln Glu Val Leu His His Ile Ser Met Ile Pro Ala Lys	
136	595 600 605	
137	tgc ctt cag gaa ggg gct gag atc tcc agt cct gca gtg gaa aga ttg	1872
138	Cys Leu Gln Glu Gly Ala Glu Ile Ser Ser Pro Ala Val Glu Arg Leu	
139	610 615 620	
140	atc agt tcg ctg ttg cga aca tgt gat ttg gaa gag ttt cag acc tgt	1920
141	Ile Ser Ser Leu Leu Arg Thr Cys Asp Leu Glu Glu Phe Gln Thr Cys	
142	625 630 635 640	
143	ctg gtg cgg cac tgc aag cat gcg ttt ggc tgt gcg ctg gtg cac acc	1968
144	Leu Val Arg His Cys Lys His Ala Phe Gly Cys Ala Leu Val His Thr	

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/434,382DATE: 11/17/1999
TIME: 13:52:25

Input Set: I434382.RAW

145	645	650	655	
146	tct ggc tgg aaa gtg gtc tat tcc	ggg gac acc atg ccc	tgc gag gct	2016
147	Ser Gly Trp Lys Val Val Tyr Ser	Gly Asp Thr Met Pro Cys Glu Ala		
148	660	665	670	
149	ctg gtc cgg atg ggg aaa gat gcc	acc ctc ctg ata cat gaa gcc acc		2064
150	Leu Val Arg Met Gly Lys Asp Ala	Thr Leu Leu Ile His Glu Ala Thr		
151	675	680	685	
152	ctg gaa gat ggt ttg gaa gag gaa	gca gtc gaa aag aca cac agc aca		2112
153	Leu Glu Asp Gly Leu Glu Glu Ala	Val Glu Lys Thr His Ser Thr		
154	690	695	700	
155	acg tcc caa gcc atc agc gtg ggg	atg cgg atg aac gcg gag ttc att		2160
156	Thr Ser Gln Ala Ile Ser Val	Gly Met Arg Met Asn Ala Glu Phe Ile		
157	705	710	715	720
158	atg ctg aac cac ttc agc cag cgc	tat gcc aag gtc ccc ctc ttc agc		2208
159	Met Leu Asn His Phe Ser Gln Arg	Tyr Ala Lys Val Pro Leu Phe Ser		
160	725	730	735	
161	ccc aac ttc agc gag aaa gtg gga	gtt gcc ttt gac cac atg aag gtc		2256
162	Pro Asn Phe Ser Glu Lys Val	Gly Val Ala Phe Asp His Met Lys Val		
163	740	745	750	
164	tgc ttt gga gac ttt cca aca atg	ccc aag ctg att ccc cca ctg aaa		2304
165	Cys Phe Gly Asp Phe Pro Thr Met	Pro Lys Leu Ile Pro Pro Leu Lys		
166	755	760	765	
167	gcc ctg ttt gct ggc gac atc	gag gag atg gag gag cgc agg gag aag		2352
168	Ala Leu Phe Ala Gly Asp Ile	Glu Glu Met Glu Glu Arg Arg Glu Lys		
169	770	775	780	
170	cgg gag ctg cgg cag gtg	cgg ggc ctc ctg tcc agg gag ctg gca		2400
171	Arg Glu Leu Arg Gln Val Arg Ala	Ala Leu Ser Arg Glu Leu Ala		
172	785	790	795	800
173	ggc ggc ctg gag gat ggg gag	cct cag aag cg gcc cac aca gag		2448
174	Gly Leu Glu Asp Gly Glu Pro Gln	Gln Lys Arg Ala His Thr Glu		
175	805	810	815	
176	gag cca cag gcc aag aag gtc	aga gcc cag tga		2481
177	Glu Pro Gln Ala Lys Lys Val Arg	Ala Gln		
178	820	825		
179	<210> SEQ ID NO 2			
180	<211> LENGTH: 826			
181	<212> TYPE: PRT			
182	<213> ORGANISM: Homo sapiens			
183	<400> SEQUENCE: 2			
184	Met Trp Ala Leu Cys Ser Leu Leu Arg Ser Ala Ala Gly Arg Thr Met			
185	1	5	10	15
186	Ser Gln Gly Arg Thr Ile Ser Gln Ala Pro Ala Arg Arg Glu Arg Pro			
187	20	25	30	
188	Arg Lys Asp Pro Leu Arg His Leu Arg Thr Arg Glu Lys Arg Gly Pro			
189	35	40	45	
190	Ser Gly Cys Ser Gly Gly Pro Asn Thr Val Tyr Leu Gln Val Val Ala			
191	50	55	60	
192	Ala Gly Ser Arg Asp Ser Gly Ala Ala Leu Tyr Val Phe Ser Glu Phe			
193	65	70	75	80
194	Asn Arg Tyr Leu Phe Asn Cys Gly Glu Gly Val Gln Arg Leu Met Gln			

PAGE: 5

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/434,382DATE: 11/17/1999
TIME: 13:52:25

Input Set: I434382.RAW

195	85	90	95
196	Glu His Lys Leu Lys Val Ala Arg Leu Asp Asn Ile Phe Leu Thr Arg		
197	100	105	110
198	Met His Trp Ser Asn Val Gly Gly Leu Ser Gly Met Ile Leu Thr Leu		
199	115	120	125
200	Lys Glu Thr Gly Leu Pro Lys Cys Val Leu Ser Gly Pro Pro Gln Leu		
201	130	135	140
202	Glu Lys Tyr Leu Glu Ala Ile Lys Ile Phe Ser Gly Pro Leu Lys Gly		
203	145	150	155
204	Ile Glu Leu Ala Val Arg Pro His Ser Ala Pro Glu Tyr Glu Asp Glu		
205	165	170	175
206	Thr Met Thr Val Tyr Gln Ile Pro Ile His Ser Glu Gln Arg Arg Gly		
207	180	185	190
208	Lys His Gln Pro Trp Gln Ser Pro Glu Arg Pro Leu Ser Arg Leu Ser		
209	195	200	205
210	Pro Glu Arg Ser Ser Asp Ser Glu Ser Asn Glu Asn Glu Pro His Leu		
211	210	215	220
212	Pro His Gly Val Ser Gln Arg Arg Gly Val Arg Asp Ser Ser Leu Val		
213	225	230	235
214	Val Ala Phe Ile Cys Lys Leu His Leu Lys Arg Gly Asn Phe Leu Val		
215	245	250	255
216	Leu Lys Ala Lys Glu Met Gly Leu Pro Val Gly Thr Ala Ala Ile Ala		
217	260	265	270
218	Pro Ile Ile Ala Ala Val Lys Asp Gly Lys Ser Ile Thr His Glu Gly		
219	275	280	285
220	Arg Glu Ile Leu Ala Glu Glu Leu Cys Thr Pro Pro Asp Pro Gly Ala		
221	290	295	300
222	Ala Phe Val Val Val Glu Cys Pro Asp Glu Ser Phe Ile Gln Pro Ile		
223	305	310	315
224	Cys Glu Asn Ala Thr Phe Gln Arg Tyr Gln Gly Lys Ala Asp Ala Pro		
225	325	330	335
226	Val Ala Leu Val Val His Met Ala Pro Ala Ser Val Leu Val Asp Ser		
227	340	345	350
228	Arg Tyr Gln Gln Trp Met Glu Arg Phe Gly Pro Asp Thr Gln His Leu		
229	355	360	365
230	Val Leu Asn Glu Asn Cys Ala Ser Val His Asn Leu Arg Ser His Lys		
231	370	375	380
232	Ile Gln Thr Gln Leu Asn Leu Ile His Pro Asp Ile Phe Pro Leu Leu		
233	385	390	395
234	400		
235	Thr Ser Phe Arg Cys Lys Lys Glu Gly Pro Thr Leu Ser Val Pro Met		
236	405	410	415
237	Val Gln Gly Glu Cys Leu Leu Lys Tyr Gln Leu Arg Pro Arg Arg Glu		
238	420	425	430
239	Trp Gln Arg Asp Ala Ile Ile Thr Cys Asn Pro Glu Glu Phe Ile Val		
240	435	440	445
241	Glu Ala Leu Gln Leu Pro Asn Phe Gln Gln Ser Val Gln Glu Tyr Arg		
242	450	455	460
243	Arg Ser Ala Gln Asp Gly Pro Ala Pro Ala Glu Lys Arg Ser Gln Tyr		
244	465	470	475
	480		
	Pro Glu Ile Ile Phe Leu Gly Thr Gly Ser Ala Ile Pro Met Lys Ile		

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.



Input Set: I434382.RAW

Line ? Error/Warning

Original Text

877 W "N" or "Xaa" used: Feature required
1109 W "N" or "Xaa" used: Feature required
1331 W "N" or "Xaa" used: Feature required

agcatca tc atgaagg a agagg t gaga tgcctgg
cagg aattca gcacata tc attgtt cagn n
cagg aattca gcacata tc attgtt cagn n